

WHAT IS CLAIMED IS:

1. A cloned DNA sequence encoding a polypeptide of hap gene, wherein the sequence has the formula

ATGTTTGACTGTATGGATGTTCTGTCACTGAGTCCTGGGCAAATCCTGATTCTACACTGCGAGTCC  
GTCTTCCTGCATGCTCCAGGAGAAAGCTCTCAAAGCATGCTTCAGTGGATTGACCCAAACCGAATG  
GCAGCATCGGCACACTGCTCAATCAATTGAAACACAGAGCACCAGCTCTGAGGAACTCGTCCCAAG  
CCCCCATCTCCACTTCCTCCCCCTCGAGTGATCAAACCCTGCTTCGTCTGCCAGGACAAATCATC  
AGGGTACCACTATGGGGTCAGCGCCTGTGAGGGATGAAGGGCTTTTTCCGCAGAAGTATTCAGAAG  
AATATGATTTACACTTGTCACCGAGATAAGAACTGTGTTATTAATAAAGTCACCAGGAATCGATGC  
CAATACTGTGACTCCAGAAGTGCTTTGAAGTGGGAATGTCCAAAGAATCTGTCAGGAATGACAGG  
AACAAGAAAAAGAAGGAGACTTCGAAGCAAGAATGCACAGAGAGCTATGAAATGACAGCTGAGTTG  
GACGATCTCACAGAGAAGATCCGAAAAGCTCACCAGGAACTTTCCTTCACTCTCGCAGCTGGGT  
AAATACACCACGAATTCCAGTGCTGACCATCGAGTCCGACTGGACCTGGGCCTCTGGGACAAATTC  
AGTGAACCTGGCCACCAAGTGCATTATTAAGATCGTGGAGTTTGCTAAACGTCTGCCTGGTTTCACT  
GGCTTGACCATCGCAGACCAAATTACCTGCTGAAGGCCGCTGCCTGGACATCCTGATTCTTAGA  
ATTTGCACCAGGTATACCCCAGAACAAAGACACCATGACTTTCTCAGACGGCCTTACCCTAAATCGA  
ACTCAGATGCACAATGCTGGATTTGGTCCTCTGACTGACCTTGTGTTACCTTTGCCAACCAGCTC  
CTGCCTTTGGAAATGGATGACACAGAAACAGGCCTTCTCAGTGCCATCTGCTTAATCTGTGGAGAC  
CGCCAGGACCTTGAGGAACCGACAAAAGTAGATAAGCTACAAGAACCATTGCTGGAAGCACTAAAA  
ATTTATATCAGAAAAAGACGACCCAGCAAGCCTCACATGTTTCCAAAGATCTTAATGAAAATCACA  
GATCTCCGTAGCATCAGTGCTAAAGGTGCAGAGCGTGTAATTACCTTGAAAATGGAAATTCCTGGA  
TCAATGCCACCTCTCATTCAAGAAATGATGGAGAATTCTGAAGGACATGAACCCTTGACCCCAAGT  
TCAAGTGGGAACACAGCAGAGCACAGTCCTAGCATCTCACCCAGCTCAGTGGAACAGTGGGGTC  
AGTCAGTCACCACTCGTGCAATAA,

Sub c  
cont.

and serotypic variants thereof, wherein said DNA is in a purified form.

2. DNA sequence as claimed in claim 1, which is free of human serum proteins, viral proteins, and nucleotide sequences encoding said proteins.

3. DNA sequence as claimed in claim 1, which is free of human tissue.

4. DNA sequence as claimed in claim <sup>57</sup>~~1~~, wherein the sequence has the formula:

GTCAGGAATGACAGGAACAAGAAAAAGAAGGAGACTTCGAAGCAAGAATGC.

5. DNA sequence as claimed in claim <sup>57</sup>~~1~~, wherein the sequence has the formula:

GCTGAGTTGGACCATCTCACAGAGAAGATCCGA.

6. DNA sequence as <sup>57</sup>~~claimed~~ in claim <sup>57</sup>~~1~~, wherein the sequence has the formula:

GGGGTCACTCAGTCACCACTCGTGCAA.

7. DNA sequence as <sup>57</sup>~~claimed~~ in claim <sup>57</sup>~~1~~, wherein the sequence has the formula:

AATGACAGGAACAAGAAAAAGAAGGAGACT.

8. DNA sequence as <sup>57</sup>~~claimed~~ in claim <sup>57</sup>~~1~~, wherein the sequence has the formula:

ATGTTTGACTGTATGGATGTTCTGTCACTGAGTCCTGGGCAAATCCTCGATTTC  
TAACTGCGAGTCCGTCTTCCTGCATGCTCCAGGAGAAAGCTCTCAAAGCATGC  
TTCAGTGGATTGACCCAAACCGAATGGCAGCATCGGCACACTGCTCAATCA.

9. DNA sequence as <sup>57</sup>~~claimed~~ in claim <sup>57</sup>~~1~~, wherein the sequence has the formula:

CATGAACCCTTGACCCCAAGTTCAAGTGGGAACACAGCAGAGCACACTCCTAGC  
ATCTCACCCAGCTGAGTGGAAAACAGTGGGGTCACTCAGTCACCACTCGTGCAA.

10. A DNA probe consisting essentially of a radionuclide bonded to the DNA sequence of claim 1.

11. A hybrid duplex molecule consisting essentially of the DNA sequence of claim 1 hydrogen bonded to a nucleotide sequence of complementary base sequence.

12. Hybrid duplex molecule as claimed in claim <sup>11</sup>9, wherein said nucleotide sequence is a DNA sequence.

13. Hybrid duplex molecule as claimed in claim <sup>11</sup>9, wherein said nucleotide sequence is a RNA sequence.

14. Hybrid duplex molecule as claimed in claim <sup>11</sup>9, wherein a radionuclide label is bonded to said DNA sequence.

15. A polypeptide comprising an amino acid sequence of hap protein, wherein the polypeptide contains the amino acid sequence

MetPheAspCysMetAspValLeuSerValSerProGlyGlnIleLeuAspPheTyrThrAla  
SerProSerSerCysMetLeuGlnGluLysAlaLeuLysAlaCysPheSerGlyLeuThrGln  
ThrGluTrpGlnHisArgHisThrAlaGlnSerIleGluThrGlnSerThrSerSerGluGlu  
LeuValProSerProProSerProLeuProProProArgValTyrLysProCysPheValCys  
GlnAspLysSerSerGlyTyrHisTyrGlyValSerAlaCysGluGlyCysLysGlyPhePhe  
ArgArgSerIleGlnLysAsnMetIleTyrThrCysHisArgAspLysAsnCysValIleAsn  
LysValThrArgAspArgCysGlnTyrCysArgLeuGlnLysCysPheGluValGlyMetSer  
LysGluSerValArgAsnAspArgAsnLysLysLysLysGluThrSerLysGlnGluCysThr  
GluSerTyrGluMetThrAlaGluLeuAspAspLeuThrGluLysIleArgLysAlaHisGln  
GluThrPheProSerLeuCysGlnLeuGlyLysTyrThrThrAsnSerSerAlaAspHisArg

ValArgLeuAspLeuGlyLeuTrpAspLysPheSerGluLeuAlaThrLysCysIleIleLys  
 IleValGluPheAlaLysArgLeuProGlyPheThrGlyLeuThrIleAlaAspGlnIleThr  
 LeuLeuLysAlaAlaCysLeuAspIleLeuIleLeuArgIleCysThrArgTyrThrProGlu  
 GlnAspThrMetThrPheSerAspGlyLeuThrLeuAsnArgThrGlnMetHisAsnAlaGly  
 PheGlyProLeuThrAspLeuValPheThrPheAlaAsnGlnLeuLeuProLeuGluMetAsp  
 AspThrGluThrGlyLeuLeuSerAlaIleCysLeuIleCysGlyAspArgGlnAspLeuGlu  
 GluProThrLysValAspLysLeuGlnGluProLeuLeuGluAlaLeuLysIleTyrIleArg  
 LysArgArgProSerLysProHisMetPheProLysIleLeuMetLysIleThrAspLeuArg  
 SerIleSerAlaLysGlyAlaGluArgValIleThrLeuLysMetGluIleProGlySerMet  
 ProProLeuIleGlnGluMetMetGluAsnSerGluGlyHisGluProLeuThrProSerSer  
 SerGlyAsnThrAlaGluHisSerProSerIleSerProSerSerValGluAsnSerGlyVal  
 SerGlnSerProLeuValGln,

and serotypic variants and fragments thereof, wherein said polypeptide is free from human serum proteins, virus, viral protein, human tissue, and human tissue components.

16. Polypeptide as claimed in claim 15, which is free from human, blood-derived protein.

17. A polypeptide as claimed in claim 15, wherein the polypeptide comprises a peptide fragment having the amino acid sequence:

GlnHisArgHisThrAlaGlnSerIleGluThrGlnSerThrSerSerGluGlu  
 LeuValProSerProProSerProLeuProProProArgValTyrLysProCysPheValCys  
 GlnAspLysSerSerGlyTyrHisTyrGlyValSerAlaCysGluGlyCysLysGlyPhePhe  
 ArgArgSerIleGlnLysAsnMetIleTyrThrCysHisArgAspLysAsnCysValIleAsn

LysValThrArgAsnArgCysGlnTyrCysArgLeuGlnLysCysPheGluValGlyMetSer  
LysGluSerValArgAsnAspArgAsnLysLysLysLysGluThrSerLysGlnGluCysThr  
GluSerTyrGluMetThrAlaGluLeuAspAspLeuThrGluLysIleArgLysAlaHisGln  
GluThrPheProSerLeuCys.

18. A polypeptide as claimed in claim 15, wherein the polypeptide comprises a peptide fragment having the amino acid sequence:

ValArgAsnAspArgAsnLysLysLysLysGluThrSerLysGlnGluCys (peptide 1);  
and serotypic variants thereof.

19. A polypeptide as claimed in claim 15, wherein the polypeptide comprises a peptide fragment having the amino acid sequence:

AsnAspArgAsnLysLysLysLysGluThrCys (peptide 2);  
and serotypic variants thereof.

20. A polypeptide as claimed in claim 15, wherein the polypeptide comprises a peptide fragment having the amino acid sequence:

CysGlyValSerGlnSerProLeuValGln (peptide 3);  
and serotypic variants thereof.

21. A polypeptide as claimed in claim 15, wherein the polypeptide comprises a peptide fragment having the amino acid sequence:

AlaGluLeuAspAspLeuThrGluLysIleArg;

and serotypic variants thereof.

22. A polypeptide as claimed in claim 15, wherein the polypeptide comprises a peptide fragment having the amino acid sequence:

MetPheAspCysMetAspValLeuSerValSerProGlyGlnIleLeuAspPheTyrThr  
AlaSerProSerSerCysMetLeuGlnGluLysAlaLeuLysAlaCysPheSerGlyLeu  
ThrGlnThrGluTrpGlnHisArgHisThrAlaGlnSer.

23. A polypeptide as claimed in claim 15, wherein the polypeptide comprises a peptide fragment having the amino acid sequence:

HisGluProLeuThrProSerSerSerGlyAsnThrAlaGluHisSerProSer  
IleSerProSerSerValGluAsnSerGlyValSerGlnSerProLeuValGln

24. A process for selecting a nucleotide sequence coding for hap protein or a portion thereof from a group of nucleotide sequences comprising the step of determining which of said nucleotide sequences hybridizes to a DNA sequence as claimed in claim 1.

25. Process as claimed in claim <sup>24</sup>/<sub>20</sub>, wherein said nucleotide sequence is a DNA sequence.

26. Process as claimed in claim <sup>25</sup>/<sub>21</sub>, wherein said nucleotide sequence is selected by Southern blot technique.

C 27. Process as claimed in claim <sup>24</sup>20, wherein said nucleotide sequence is a RNA sequence.

C 28. Process as claimed in claim <sup>27</sup>23, wherein said nucleotide sequence is selected by Northern blot technique.

C 29. Process as claimed in claim <sup>24</sup>20, wherein said process comprises the step of detecting a label bonded to said DNA sequence.

C 30. Process as claimed in claim <sup>29</sup>25, wherein said label is a radionuclide.

31. A recombinant vector comprising lambda-NM1149 having an EcoRI restriction endonuclease site into which has been inserted the DNA sequence as claimed in claim 1.

32. Plasmid pCOD20.

C 33. An E. coli bacterial culture in a purified form, wherein the culture comprises E. coli cells containing <sup>a plasmid</sup>DNA, wherein a portion of said <sup>plasmid</sup>DNA comprises the DNA sequence as claimed in claim 1.

C 34. Bacterial culture as claimed in claim 27, wherein said cells are comprised of E. coli strain TG-1.

35. A method for assaying a fluid for the presence of an agonist or antagonist to retinoic acid receptor RAR- $\beta$ , wherein the method comprises

(A) providing an aqueous solution containing a known concentration of the proteinaceous receptor as claimed in claim 13;

(B) incubating the receptor with the fluid suspected of containing the agonist or antagonist under conditions sufficient to bind the receptor to the agonist or antagonist; and

(C) determining whether there is change in concentration of the proteinaceous receptor in the aqueous solution.

36. Method as claimed in claim 35, wherein the receptor and the agonist or antagonist form a complex.

37. Method as claimed in claim 36, wherein a crosslinking agent is present in an amount sufficient to inhibit dissociation of the receptor and the agonist or antagonist.

38. A cloned DNA sequence encoding a polypeptide of hap gene, wherein the sequence has the formula

CCCATGC

GAGCTGTTTGAAGGACTGGGATGCCGAGAACGCGAGCGATCCGAGCAGGGTTTGTCTGGGCACCGT  
ATGTTTGACTGTATGGATGTTCTGTCACTGAGTCCTGGGCAAATCCTGATTCTACACTGCGAGTCC  
GTCTTCCTGCATGCTCCAGGAGAAAGCTCTCAAAGCATGCTTCAGTGGATTGACCCAAACCGAATG  
GCAGCATCGGCACACTGCTCAATCAATTGAAACACAGAGCACCAGCTCTGAGGAACTCGTCCCAAG  
CCCCCATCTCCACTTCCTCCCCCTCGAGTGATCAAACCCTGCTTCGTCTGCCAGGACAAATCATC  
AGGGTACCACTATGGGGTCAGCGCCTGTGAGGATGAAGGGCTTTTTCCGCAGAAGTATTCAGAAG  
AATATGATTTACACTTGTCACCGAGATAAGAAGTGTGTTATTAATAAAGTCACCAGGAATCGATGC  
CAATACTGTCTGACTCCAGAAGTGCTTTGAAGTGGGAATGTCCAAAGAATCTGTCAGGAATGACAGG  
AACAAGAAAAAGAAGGAGACTTCGAAGCAAGAATGCACAGAGAGCTATGAAATGACAGCTGAGTTG  
GACGATCTCACAGAGAAGATCCGAAAAGCTCACCAGGAACTTTCCCTTCACTCTCGCAGCTGGGT  
AAATACACCACGAATTCAGTGCTGACCATCGAGTCCGACTGGACCTGGGCCTCTGGGACAAATTC  
AGTGAAGTGGCCACCAAGTGCATTATTAAGATCGTGGAGTTTGCTAAACGTCTGCCTGGTTTCACT  
GGCTTGACCATCGCAGACCAAATTACCCTGCTGAAGGCCGCTGCCTGGACATCCTGATTCTTAGA  
ATTTGCACCAGGTATACCCAGAACAAAGACACCATGACTTTCTCAGACGGCCTTACCCTAAATCGA  
ACTCAGATGCACAATGCTGGATTTGGTCCTCTGACTGACCTTGTTTACCTTTGCCAACCAGCTC



CTGCCTTTGGAAATGGATGACACAGAAACAGGCCTTCTCAGTGCCATCTGCTTAATCTGTGGAGAC  
CGCCAGGACCTTGAGGAACCCACAAAAGTAGATAAGCTACAAGAACCATTGCTGGAAGCACTAAAA  
ATTTATATCAGAAAAAGACGACCCAGCAAGCCTCACATGTTTCCAAAGATCTTAATGAAAATCACA  
GATCTCCGTAGCATCAGTGCTAAAGGTGCAGAGCGTGTAATTACCTTGAAAATGGAAATTCCTGGA  
TCAATGCCACCTCTCATTCAAGAAATGATGGAGAATTCTGAAGGACATGAACCCTTGACCCCAAGT  
TCAAGTGGGAACACAGCAGAGCACAGTCCCTAGCATCTCACCCAGCTCAGTGGAAAACAGTGGGGTC  
AGTCAGTCACCACTCGTGCAATAA,

and serotypic variants thereof, wherein said DNA is in a purified form.

39. <sup>a</sup> DNA sequence as claimed in claim <sup>59</sup>~~38~~, which is free of human serum proteins, viral proteins, and nucleotide sequences encoding said proteins.

40. <sup>a</sup> DNA sequence as claimed in claim 1, which is free of human tissue.

41. A DNA probe consisting essentially of a radionuclide bonded to the DNA sequence of claim <sup>59</sup>~~38~~.

42. A hybrid duplex molecule consisting essentially of the DNA sequence of claim <sup>59</sup>~~38~~ hydrogen bonded to a nucleotide sequence of complementary base sequence.

43. Hybrid duplex molecule as claimed in claim 11, wherein said nucleotide sequence is a DNA sequence.

44. Hybrid duplex molecule as claimed in claim 11, wherein said nucleotide sequence is a RNA sequence.

45. Hybrid duplex molecule as claimed in claim 11, wherein a radionuclide label is bonded to said DNA sequence.

Sub  
C3  
46. A process for selecting a nucleotide sequence coding for hap protein or a portion thereof from a group of nucleotide sequences comprising the step of determining which of said nucleotide sequences hybridizes to a DNA sequence as claimed in claim 38.

C  
47. Process as claimed in claim <sup>46</sup>~~24~~, wherein said nucleotide sequence is a DNA sequence.

C  
48. Process as claimed in claim <sup>47</sup>~~24~~, wherein said nucleotide sequence is selected by Southern blot technique.

49. Process as claimed in claim 46, wherein said nucleotide sequence is a RNA sequence.

C  
50. Process as claimed in claim <sup>48</sup>~~46~~, wherein said nucleotide sequence is selected by Northern blot technique.

51. Process as claimed in claim 46, wherein said process comprises the step of detecting a label bonded to said DNA sequence.

52. Process as claimed in claim 51, wherein said label is a radionuclide.

53. A recombinant DNA molecule comprising a DNA sequence of coding for a retinoic acid receptor, said DNA sequence coding on expression in a unicellular host for a polypeptide displaying the retinoic acid and DNA binding properties of RAR-β and being operatively linked to an expression control sequence in said DNA molecule. ✓

54. Plasmid pPROHAP. ✓

55. An E. coli bacterial culture in a purified form, where-  
in the culture comprises E. coli cells containing <sup>a plasmid</sup> ~~DNA~~, wherein a  
portion of said <sup>plasmid</sup> ~~DNA~~ comprises the DNA sequence as claimed in  
claim <sup>57</sup> ~~38~~.

56. Bacterial culture as claimed in claim 55, wherein said  
cells are comprised of E. coli strain DH5 $\alpha$ F'.

add c<sup>4</sup>

add  
D<sup>1</sup>

add  
e<sup>14</sup>

add  
G<sup>1</sup>